Received: 17 August 2020 | Revised: 3 December 2020 | Accepted: 21 January 2021

DOI: 10.1002/rob.22016

FIELD REPORT

WILEY

The MADMAX data set for visual-inertial rover navigation on Mars

Lukas Meyer[1] | Michal Smíšek[1] | Alejandro Fontan Villacampa[1] | Laura Oliva Maza[1] | Daniel Medina[2] | Martin J. Schuster[1] | Florian Steidle[1] | Mallikarjuna Vayugundla[1] | Marcus G. Müller[1] | Bernhard Rebele[3] | Armin Wedler[1] | Rudolph Triebel[1]

[1]Department of Perception and Cognition, Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany

[2]Department of Nautical Systems, Institute of Communications and Navigation, German Aerospace Center (DLR), Neustrelitz, Germany

[3]Analysis and Control of Advanced Robotic Systems, Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany

Correspondence

Lukas Meyer, Department of Perception and Cognition, Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Muenchener Strasse 20, 82234 Wessling, Germany.
Email: lukas.meyer@dlr.de

Funding information

MOdulares Robotisches EXplorationssystem (MOREX)

Abstract

Planetary rovers increasingly rely on vision-based components for autonomous navigation and mapping. Developing and testing these components requires representative optical conditions, which can be achieved by either field testing at planetary analog sites on Earth or using prerecorded data sets from such locations. However, the availability of representative data is scarce and field testing in planetary analog sites requires a substantial financial investment and logistical overhead, and it entails the risk of damaging complex robotic systems. To address these issues, we use our compact human-portable DLR Sensor Unit for Planetary Exploration Rovers (SUPER) in the Moroccan desert to show resource-efficient field testing and make the resulting Morocco-Acquired data set of Mars-Analog eXploration (MADMAX) publicly accessible. The data set consists of 36 different navigation experiments, captured at eight Mars analog sites of widely varying environmental conditions. Its longest trajectory covers 1.5 km and the combined trajectory length is 9.2 km. The data set contains time-stamped recordings from monochrome stereo cameras, a color camera, omnidirectional cameras in stereo configuration, and from an inertial measurement unit. Additionally, we provide the ground truth in position and orientation together with the associated uncertainties, obtained by a real-time kinematic-based algorithm that fuses the global navigation satellite system data of two body antennas. Finally, we run two state-of-the-art navigation algorithms, ORB-SLAM2 and VINS-mono, on our data to evaluate their accuracy and to provide a baseline, which can be used as a performance reference of accuracy and robustness for other navigation algorithms. The data set can be accessed at https://rmc.dlr.de/morocco2018.

KEYWORDS

exploration, extreme environments, navigation, planetary robotics, SLAM

1 | INTRODUCTION

Planetary surfaces are mostly explored by mobile robotic platforms, because these environments are difficult and expensive to reach for humans and exhibit hazardous environmental conditions. Owing to the interrupted and high-latency communication, as well as the lack of prior knowledge about the environment, such mobile robotic platforms have to operate autonomously to some extent.

The quality of autonomous decision-making by a mobile robot depends heavily on the navigation and mapping software solutions, as precise localization capabilities are crucial for experiments to be completed successfully. The navigation and mapping solutions increasingly rely on vision-based components, as camera systems are commonly used in space missions since they are robust in harsh environments and space-qualified systems already exist. Extraterrestrial vision-based robot navigation needs to be able to operate under optical conditions that differ greatly from most locations on Earth. While a software simulation can bootstrap initial phases of component development and testing, to satisfy the robustness requirements it must usually be tested against inputs with representative optical conditions. The maturity and usability of such navigation solutions therefore depends on the availability of the sensor input representative of the targeted planetary environment for development and testing. These data are typically obtained on Earth in planetary analog locations with optical features similar to the targeted extraterrestrial bodies.

However, planetary analog sites on Earth are generally remote locations (see Preston et al., 2012, for an overview) that are difficult to access. Bringing a robotic system to such areas for onsite testing usually results in costly logistics, comes with high demand in operations personnel, and involves the risk of damaging the robotic system. Offsite testing using previously recorded data sets is a more cost-efficient approach, but suffers from the scarcity of data sets available for this purpose. Furthermore, data sets allow software components to be run repeatedly to improve algorithms and test them more efficiently. Algorithms can also be compared with each other with respect to the accuracy, robustness, and computational performance.

This paper addresses the onsite and offsite testing of vision-based navigation solutions in planetary analog scenarios by considering three aspects:

- We discuss the use of hand-held sensor devices that resemble the sensor setup of a planetary rover as a cost-efficient, low-risk alternative to onsite field testing. For this, we present the human-portable Sensor Unit for Planetary Exploration Rovers (SUPER) of the German Aerospace Center (DLR) and its application in the 2018 Mars-analog field test in the Moroccan desert.
- During this field test, we recorded the Morocco-Acquired Data set of Mars-Analog eXploration (MADMAX). It is a comprehensive collection of visual inertial navigation data representative of a Mars-roving scenario, containing 36 trajectories with a combined length of 9.2 km. We describe the data set in detail and make it publicly available.
- We use MADMAX with two state-of-the-art navigation algorithms, ORB-SLAM2 (Mur-Artal & Tardós, 2017) and VINS-Mono (Qin et al., 2018), to evaluate their algorithmic performance and to evaluate the challenges of the data set. We show that the hand-held MADMAX can be considered representative for planetary rover navigation by comparing the navigation results in terms of accuracy with results obtained from a planetary rover prototype. Additionally, we use the results to provide a navigation baseline for our Mars-analog scenario that can be used by other navigation algorithms as a performance reference for accuracy and robustness.

1.1 | Hand-held field testing

Testing of planetary rovers and their navigation algorithms can be performed with increasing complexity levels. Tests in laboratory environments and in artificially created outdoor testbeds are a good initial way to validate navigation and mapping solutions. A multitude of rover navigation solutions have been tested in such scenarios, for example,

- ExoMars Test Rover (ExoTeR) is used to test localization and mapping components both in an indoor laboratory environment and in an outdoor test-facility of the European Space Agency (Hidalgo-Carrió et al., 2018).
- Rovers Minnie and Mana are used outdoors on artificially created terrain that resembles the features of Mars (Post et al., 2018).

For a state-of-the-art validation of planetary rovers in general, and especially for navigation solutions, field tests in analog environments become necessary. These analog environments are usually remote regions like deserts or volcanoes that resemble the environments of different celestial bodies, typically the Moon or Mars. A detailed list of analog sites on Earth is provided by Preston et al. (2012). In the past, many different planetary rover prototypes were placed in such environments for system testing or testing of the full rover mission including scientific operations. The list of such endeavors known to us comprises:

- The long-distance rover traverses in the Atacama Desert in Chile, pioneered by Wettergreen et al. (1999) and Wettergreen et al. (2005) with teleoperation and partial autonomy for the rovers Nomad and Zoë, respectively. And more recently, the long-range autonomous exploration tests by the Seeker rover (Woods et al., 2014) in the same area.
- Tests for the ExoMars rover mission, such as the SAFER field test (Gunes-Lasnet et al., 2014) in the Atacama Desert and the ExoFit rover tests (Motaghian et al., 2019) in southern Spain.
- The MARS2013 mission in Morocco, where full Mars-analog operations were tested, including scientists in the field, communication infrastructure, and rovers (Groemer et al., 2014).
- The Utah field trials in the United States testing the SherpaTT rover and Coyote III robot, with a the focus on multirobot systems

and teleoperation (Sonsalla et al., 2017) and locomotion capabilities (Cordes et al., 2018).

- The Mojave Desert field test in the United States (Bakambu et al., 2012) with a focus on evaluating navigation algorithms in a Mars-analog environment, specifically considering the visual motion estimation and inertial measurement unit (IMU) enhanced wheel odometry.
- A field test in the Teide Volcano National Park on the island of Tenerife, Spain, the results of which are used, among others, by Geromichalos et al. (2020) to validate a Simultaneous Localization and Mapping (SLAM) solution.
- Visual odometry with omnidirectional cameras is evaluated with data from a field test in the Atacama Desert by Corke et al. (2004).
- Our ROBEX demo mission on top of the volcano Etna in Italy focusing on a full modular Moon-analog mission (Wedler et al., 2017), done by the DLR in 2017.

All these tests campaigns were either a full rover mission, or even a full scientific mission scenario. Generally, planetary rovers are highly integrated mechatronic systems, which makes field testing complex, something we experienced ourselves during the ROBEX demo mission (Wedler et al., 2017). The complexity is due to the multitude of components used, which require the presence of many specialists and equipment at the test site. This results in costly logistics and high manpower requirements. Furthermore, many of these endeavors are used to test various system components of the rover, allocating only limited time for navigation tests and possibly running the risk of technical failures that may delay or endanger the field test.

Our paper's first contribution is to address the complexity of field tests by providing a simplified hand-held rover navigation platform that focuses solely on the visual inertial navigation systems (VINS) components of a planetary rover. This is achieved by our SUPER hand-held device, which resembles a planetary rover in terms of sensors but leaves out all other components, such as locomotion, scientific instruments, and representative communication concepts. As SUPER is not a full planetary rover prototype like the LRU, the representativeness for a rover operations scenario on a celestial body can be questioned. However, our analysis in Section 6.3 suggests that the hand-held approach is representative for navigation. We therefore consider this approach as the optimal trade-off between costs and representativeness. A similar approach is taken by Furgale et al. (2012), where a pushcart platform equipped with stereo cameras, a sun sensor, and inclinometers is used for a combined Mars and Moon-analog navigation experiment on Devon Island, in the Arctic North of Canada.

We use SUPER in a Mars-analog site and perform navigation experiments to show that this kind of sensor unit allows for resource-efficient field testing thanks to three factors: its small size, the use of key hardware components only, and the fact that only two persons are needed to operate the device. In our case, the Mars-analog site is located in the north-western region of the Sahara Desert, close to the city of Erfoud in the Drâa-Tafilalet region of Morocco, as shown in Figure 2.

1.2 | Related field test data sets

Only a few publicly available vision-based navigation and mapping data sets exists, which specifically target planetary robotics. One is the Katwijk Beach Planetary Rover Data set (Hewitt et al., 2018), where a planetary rover prototype performs several long-range traverses on a beach using stereo cameras and Lidar. Another is the resulting data set from the previously mentioned experiments on Devon Island, where the pushcart platform was used to record a long trajectory of 10 km (Furgale et al., 2012). We recorded two long range navigation runs on the outskirts of the volcano Etna during the ROBEX campaign, using a lightweight planetary exploration rover prototype that used stereo vision, IMU data, and wheel odometry for navigation (Vayugundla et al., 2018). Lamarre et al. (2020) present recordings from a Mars-analog outdoor laboratory run by the Canadian Space Agency that include data for visual inertial and omnidirectional camera-based navigation. They also consider energy-budget-aware navigation by providing solar irradiation data. Lacroix et al. (2019) present navigation data recorded by several cameras, fiber optic gyro, IMU, and Lidar from the two rovers Minnie and Mana. These were obtained in the Moroccan desert—in the same region as MADMAX. Finally, the Planetary Data System (NASA, 2019) from NASA makes it possible to access data directly obtained from current and past Mars missions.

As we deployed SUPER in the Moroccan desert, we use it to record a comprehensive visual inertial navigation data set, which constitutes the second contribution of this paper. The small size and mobility of SUPER allows it to cover several different locations with varying terrains and record a variety of trajectories in a short time. In addition, we use the mobility of SUPER to record data in sites that are not yet accessible for the current generation of wheeled planetary rovers, owing to the harsh terrain. In total, we collect data from 36 experiments at eight different locations in three general areas, each with an individual geological character. The recorded sensor input consists of a monochrome stereo camera pair, a RGB color camera, two omnidirectional cameras mounted in a vertical stereo setup, and an IMU. We also compute the ground truth pose from a Real-time kinematic (RTK) Global Navigation Satellite System (GNSS) with two antennas mounted on SUPER, which allows us to obtain a ground truth not only in the position but also in the orientation.

MADMAX lies in line with the Etna data set (Vayugundla et al., 2018) as we used closely related systems with similar sensors and an identical software infrastructure in both cases. With these two data sets, it becomes possible to evaluate VINS algorithms in Moon- and Mars-analog scenarios at the same time, without having to adapt to a different system setup.

The Morocco data from Lacroix et al. (2019) combined with MADMAX allow to evaluate vision based navigation algorithms in similar environments—sometimes even in the same location—on very different systems. One of the main differences is the use of a longer stereo baseline for their cameras (270 mm vs. our 90 mm) that makes it possible to consider more distant features for navigation, whereas our configuration focuses on nearby scenery and local mapping. Lacroix et al. (2019) additionally records Lidar data, whereas SUPER

provides omnidirectional stereo images as secondary sensor data. The wheeled platforms of Minnie and Mana provide a rover-like movement of the system. We exploit the higher mobility of SUPER to access rougher terrain to collect data. These data become relevant for next-generation planetary rovers with improved locomotion capabilities. Lacroix et al. (2019) provide data where trajectories were traversed several times by the rovers; we instead cover more trajectories of varied character in each location. We provide both a five and a six degrees of freedom (DoF) ground truth compared to the three DoF ground truth that is normally included in planetary navigation data sets. Finally, along with our data set we include an evaluation using two state-of-the-art navigation algorithms.

The three mentioned data sets therefore allow the development and evaluation of robust navigation algorithms that are able to perform independently of system architecture or environment.

1.3 | Navigation algorithm performance reference

We use MADMAX with two state-of-the-art SLAM-based navigation algorithms, the visual odometry algorithm ORB-SLAM2 (Mur-Artal & Tardós, 2017), and the visual inertial odometry algorithm VINS-Mono (Qin et al., 2018) to evaluate their performance in a Mars-analog scenario as the final contribution of our paper. The variety of MADMAX enables us to test the navigation solutions for optimal scenarios but also challenging corner cases.

We compare the navigation accuracy of this hand-held data set to the results of navigation sequences from an additional SUPER system, this time attached to a rover. This second sensor unit was integrated with the planetary rover prototype illustrated in the background of Figure 1, and was used to record several navigation sequences. We apply identical evaluation methods in both cases to study potential differences on the navigation performance. In the end, this experiment allows us to emphasize that MADMAX can be considered as representative for planetary rover navigation.

FIGURE 1 The two SUPER units in the Moroccan desert on the *Rissani 1* location: One unit is mounted on the SherpaTT rover (Cordes et al., 2018) of the DFKI Robotics Innovation Center (background) and the other is used as human-carried device (foreground). The data presented in this article were captured by the hand-held device [Color figure can be viewed at wileyonlinelibrary.com]

In addition, the results from the state of the art can be used as a baseline for other navigation algorithms. Our approach is similar to that of Antonini et al. (2020), however in our case targeted at planetary robotics instead of indoor unmanned aerial vehicle operation. To the best of our knowledge, no such evaluation and publicly available state-of-the-art performance reference for navigation is available for planetary rover navigation yet.

1.4 | Outline

The article is structured as follows. In Section 2, we describe the field test scenario in the Moroccan desert, where we recorded MADMAX. We present our sensor suite SUPER in Section 3, outline the system specifications, provide details on the installed sensors, and describe the reference frame definitions. We introduce the experiment setup in Section 4 together with operational aspects of field testing, our approach for dual-antenna RTK GNSS ground truth computation, and the sensor calibration. We present the resulting data set with its 36 trajectories in Section 5 and discuss the specific characteristics of the different experiment locations and of the individual trajectories. We provide a detailed overview of all sensor data that can be found in the data set and how the different sensor readings can be related to each other spatially and temporally. We also address challenges that come up in MADMAX, such as influences from the extreme environment, complications that we faced during the testing, as well as lessons learned from the operation. Finally, in Section 6, we discuss the results of two state-of-the-art navigation algorithms with MADMAX and provide a performance analysis of them.

2 | SCENARIO OVERVIEW AND LOCATION

SUPER is designed according to the specifications by the European Space Robotics Technologies Research Cluster (SRC) program to test perception algorithms for planetary exploration. It is conceptualized in a versatile way, such that it can be either carried by a human in the stand-alone fashion or integrated with a robotic system via mechanical, electrical, and data interfaces. Two SUPER units were used in both configurations as shown in Figure 1. They were deployed in a 40-week field test in the Moroccan desert, close to the city of Erfoud, located in the northern Sahara during November and December 2018. Many areas in this region resemble an ideal Mars-analog in terms of optical conditions (Preston et al., 2012). The experiment site locations are marked in Figure 2 and impressions from the sites are shown in Figures 1 and 3.

There, we utilized the first SUPER in a hand-held approach to record MADMAX. The second SUPER was used in the same area for the final validation experiments of the SRC technology development roadmap

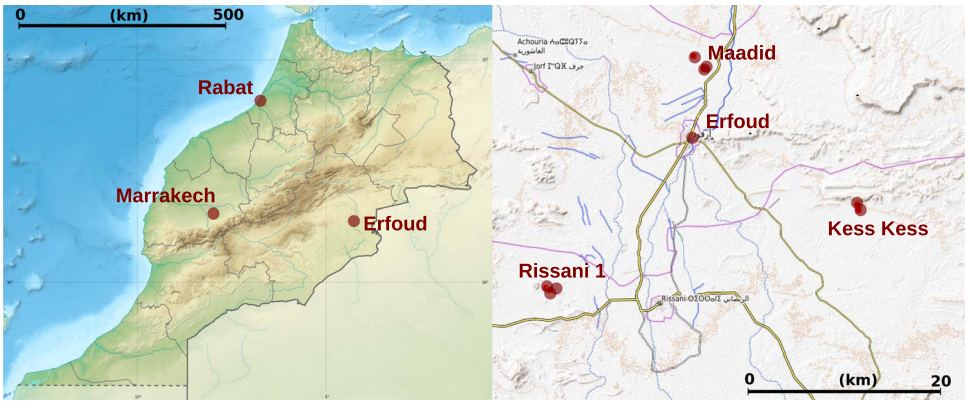


FIGURE 2 Overview of the experiment locations. All experiments were performed in the vicinity of the city of Erfoud in the northern Sahara region. Map of Morocco (left) by Eric Gaba, local map (right) by OpenStreetMap-Contributors and OpenTopoMap, both distributed under a CC BY-SA 3.0 license [Color figure can be viewed at wileyonlinelibrary.com]

PERASPERA with the projects *InFuse* and *Facilitators* as a provider of sensoric data for localization, environment mapping, environment reconstruction, and visual tracking (Post et al., 2018).[1] For this purpose, it was integrated with the SherpaTT rover (Cordes et al., 2018) as shown in the background of Figure 1. This paper focuses on the field-testing with the hand-held SUPER and publishes only the data from the hand-held experiments. For information on all other experiments, see Brinkmann et al. (2019). Nevertheless, navigation data from the rover-mounted SUPER is used to show that MADMAX is representative for planetary rover navigation, and the evaluation results are included here.

Additional vision-based navigation data resulting from InFuse and the Morocco Field Test is described in Lacroix et al. (2019), where the rovers Minnie and Mana perform navigation experiments in several desert areas around the city of Erfoud. As mentioned before, MADMAX, the data set presented here, and the data set captured by Minnie and Mana can be seen as complementary in nature.

The data presented in MADMAX is widely varied, as the small size and mobility of SUPER allows it to access several different locations for experimentation in a relatively short time. The region around Erfoud offers a rich variety of terrains: from flat to hilly, from sandy and featureless through pebbly to rocky with features of high saliency, from horizon landmarks being virtually nonexistent to salient landmarks on the horizon, from easily traversable areas to slopes nontraversable by locomotion systems of current planetary robots—such as high-inclination hillsides or sandy dune fields.

3 | SYSTEM OVERVIEW

The design of SUPER is inspired by the DLR mobile robotic systems, lightweight rover unit (LRU), a four-wheeled full-body-actuated planetary rover prototype, and ARDEA, a microaerial vehicle (MAV).

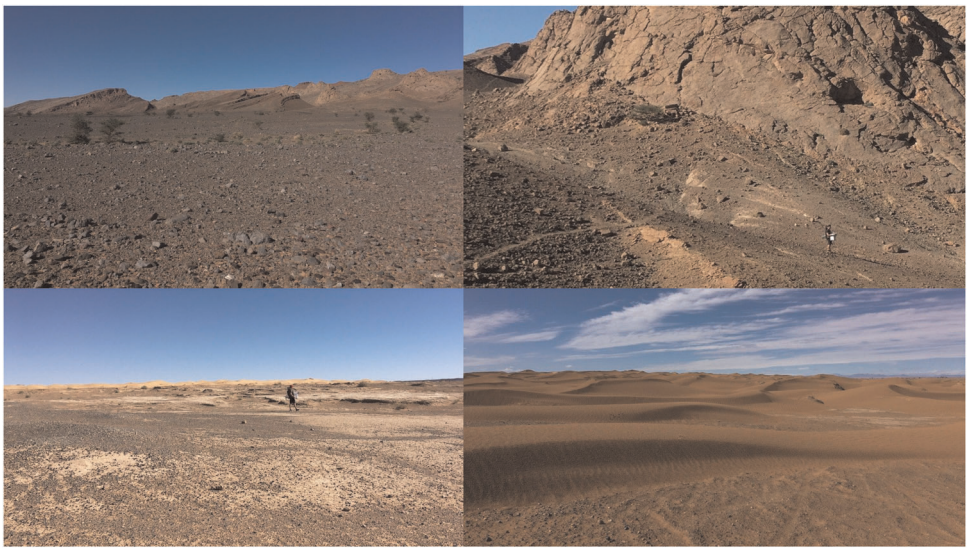


FIGURE 3 Impressions from the experiment locations: *Kess Kess* with the locations D (top left) and E (top right) and *Maadid* with locations F (bottom left) and H (bottom right) [Color figure can be viewed at wileyonlinelibrary.com]

Both systems were developed for the scenario of planetary exploration, share the same software architecture with SUPER, and use similar algorithms. See Schuster et al. (2019), for details on the LRU system and Lutz et al. (2020) for a comprehensive description of ARDEA.

The core sensor and processing components of SUPER are a copy of the components used on LRU. LRU's navigation, locomotion, and manipulation capabilities have been manifested in multiple campaigns, including the SpaceBotCamp challenge in 2015 (Schuster et al., 2019) and the ROBEX demo mission, a month-long Moon-analog field test on Mt. Etna in Sicily in 2017 (Wedler et al., 2017).

While such competitions and field tests provide good opportunities for system testing and data recording, they come at significant organizational and logistics costs. To minimize costs while maximizing the scientific yield, the design of SUPER is focused solely on planetary rover sensors. The result is a device almost identical to LRU in perception, on-board data processing, and power management capabilities, but one that omits other aspects such as active locomotion and manipulation. See Figure 4, for a detailed view.

As it was already mentioned, SUPER can be used in two scenarios as is shown in Figure 1: it can be either carried by a human or mounted on a robotic platform and integrated with that robot. Depending on the scenario, SUPER is either powered by two batteries with a capacity of 208 Wh each and hot-swap capabilities during stand-alone operations, or can be powered by a carrier robotic platform. The batteries allow it to be operated for up to 5 h. It weighs 14 kg and has a core-body size of $34 \times 26 \times 40$ cm.[2] Two on-board computers are integrated, both Kontron mITX-KBL boards with CoreTM i7-7820EQ CPUs and one Xilinx Spartan-6 LX75T FPGA for depth image computation.

Optionally, two GNSS antennas can be mounted to the sides with a wingspan of 1.28 m. We make use of the two antennas to provide a five DoF ground truth that includes not only the position but also roll and yaw information. Together with the antennas on the body of SUPER, a base GNSS station is installed on each experiment location to eliminate atmospheric delays, thus allowing precise positioning estimates. The computation of the ground truth is discussed in Section 4.3.

SUPER is focused on the perception aspect of planetary robotic applications. To keep the system simple, actuators were excluded from the design. This design choice implies that SUPER does not possess an active pan-tilt unit to change the camera orientation. The cameras point downwards at a fixed pitch of 28° relative to the body of SUPER. The camera orientation can be actively guided by the carrier— especially its heading angle—and is aligned with the orientation of the carrier. The height of the sensors depends on the height of the carrier and the adjustments made to the harness that is used by the carrier. Generally, during our hand-held experiments, the stereo camera bench is located approximately 1.20 m above the ground.

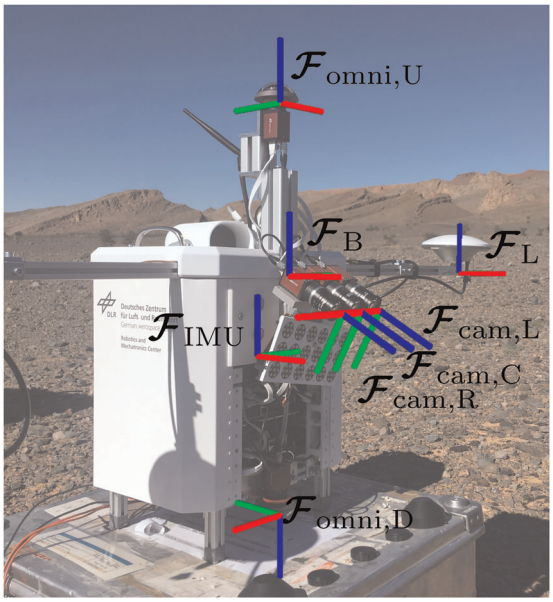


FIGURE 4 The SUPER sensor setup with the most relevant reference frames (color-scheme: *X*—red, *Y*—green, and *Z*—blue). The body frame (B), the IMU frame, the frames of the upwards (omni,U) and downwards (omni,D) facing omnidirectional cameras, the front-facing left (cam,L), right (cam,R), color (cam,C) cameras, and the left GNSS antenna (L) are shown. IMU, inertial measuring unit [Color figure can be viewed at wileyonlinelibrary.com]

Furthermore, the hand-held approach implies the absence of wheels, therefore no wheel odometry is available—a sensor input typically present in planetary robotic platforms. The data from SUPER is only targeted at VINS algorithms for navigation and mapping, which can be developed independently from the wheel odometry or other sensor inputs. Wheel odometry is a challenging scientific topic by itself and was omitted for this field test. Interested readers can learn more about our investigation into this topic in Bussmann et al. (2018), where slip of the LRU was investigated on a Moon-analog site.

3.1 | SUPER as stereo and VINS

An overview of the sensors of SUPER is given in Table 1. The placement of the sensors together with the relevant coordinate frames is shown in Figure 4.

SUPER is equipped with an optical bench carrying three cameras mounted in a row at the front of the device, with parallel optical axes. The left and right cameras are monochrome and set apart at a 90 mm baseline. These constitute the stereo camera bench, which is our primary navigation sensor. The color camera is mounted centrally between them. Data from the color camera can be used as an additional navigation source, or for finding landmarks and points of

[1]In the context of InFuse, SUPER is referred to as *Hand-held Central Rover Unit*.

[2]SUPER's dimensions are the result of using identical components as the LRU. This is desired for comparability between the systems, however contributes to a nonoptimal design in terms of size and weight.

TABLE 1 Main components of SUPER that were used for recording MADMAX

Sensor	Name	Specifications
Navigation cameras	AlliedVision Mako G-319	14 Hz, monochrome 1032 × 772 px images, rectified, auto exposure
RGB color camera	AlliedVision Mako G-319	4 Hz, color 2064 × 1544 px images, rectified, auto exposure
Camera lenses	RICOH FL-HC0614-2M	6 mm, F/1.4
Omnicam cameras	AlliedVision Mako G-319	4–8 Hz, monochrome 2064 × 1544 px images, auto exposure
Omnicam lens	Entaniya 280 Fisheye	1.07 mm, F/2.8, 280° × 360° field of view
IMU	XSENS MTi-10	100 Hz, three-axis acceleration and three-axis angular rates
GNSS receiver	Piksi Multi GNSS SwiftNav	1 Hz, GNSS Data
GNSS antenna	SwiftNav GPS500	Frequencies GPS L1/L2, GLONASS L1/L2 and BeiDou B1/B2/B3

Note: They are listed with detailed specifications and, if applicable, the sensor frequency.
Abbreviations: GNSS, global navigation satellite system; IMU, inertial measuring unit.

interest. The three-axis accelerations and angular velocities are measured by a MEMS IMU.

The exposure mode for all cameras is set to automatic with a target mean histogram level of the image at 40% brightness, which results on average exposure times of 0.01–0.001 s for the stereo camera pair and 0.02–0.005 s for the color camera. All cameras have a horizontal field of view of 61° each. This field of view, together with the stereo baseline of 90 mm, allows for a stereo overlap starting at a minimum distance of 75 mm.

The processing pipeline for image acquisition, image rectification, and depth image computation is identical to that of the LRU (Schuster et al., 2019). We use the Semi-Global Matching algorithm (Hirschmüller, 2008) to compute the depth images online onboard. This depth image stream is considered to be an intermediary data product and is included in the MADMAX data set. Note that the depth image computation is adjusted to the relevant working distance, that is, it considers a maximum disparity of 128 px, which relates to a minimum depth of 60 cm.

3.2 | Omnidirectional navigation

The configuration of SUPER is easy to modify thanks to its design, which provides mechanical, electrical, and data connections for adding extra components. We use this advantage to include additional perception sensors in our experiments. Inspired by the wide field of view from ARDEA (Lutz et al., 2020), we add a long-baseline omnidirectional camera stereo bench to the system. Omnidirectional cameras are a promising addition or alternative to vision-based navigation for planetary exploration, thanks to their high field of view. They are becoming popular for navigation and are considered in the field of planetary robotics by, for example, Corke et al. (2004) and Lamarre et al. (2020).

Our cameras have a field of view of 360° horizontally and 280° vertically and they are mounted as vertical stereo setup, in a coaxial configuration with a baseline of 61 cm. They are oriented in such a way that the upper camera has the optical axis pointing upwards and the lower camera has its optical axis pointing downwards. Their exposure mode is set to automatic with a target mean histogram level of the image at 60% brightness for the upward-facing camera, and at 30% brightness for the downward-facing camera. The different target levels are set to compensate for the bright sky or the shadows and terrain, respectively.

Impressions from the stereo setup of both cameras are shown in Figure 5. We include the recorded data streams for this sensor pair in MADMAX for all trajectories, together with the associated calibration information.

Developing and testing omnidirectional navigation solutions on SUPER was beyond the scope of the Morocco 2018 campaign and for this publication. However, it is a promising field of research and will form the subject of for future work. We hereby invite interested readers to use this collection of data for the development of novel navigation means for planetary robots.

There are multiple ways in which these omnidirectional cameras can contribute to the overall navigation solution. First, the relatively large baseline and a large overlap in the fields of view of both cameras (illustrated in Figure 5) makes it possible to formulate an omnidirectional stereo-based visual odometry. Second, both cameras

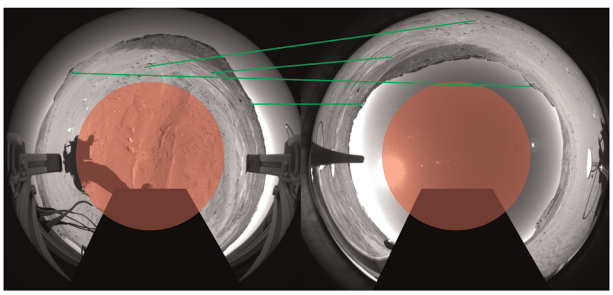


FIGURE 5 Omnidirectional stereo images from a plain experiment site with a hill range in the background. The area without stereo-overlap (red) and exemplary feature correspondences (green) between the two images are shown. The image area occupied by the operator is blacked out [Color figure can be viewed at wileyonlinelibrary.com]

have the potential to see distant landmarks on the horizon and use them for absolute orientational localization. Third, a Sun-tracking solution can be formulated for the upper camera for the absolute orientation. Fourth, the lower camera can use the optical flow of features on the ground for improved navigation. In addition, on a full rover system, the lower camera can be used as a tool for visual inspection of the locomotion system's health. These are a few suggested uses of the omnidirectional stereo data for navigation and hazard-avoidance purposes.

3.3 | Reference frames

The relevant reference frames of SUPER are annotated in Figure 4 and are listed in Table 2. The resulting transformations between the frames are included in the data set. The IMU data is referenced to the $\mathcal{F}_{\text{IMU}}$ frame. The stereo cameras provide image data with respect to the frames of the monochrome cameras $\mathcal{F}_{\text{cam,L}}$ and $\mathcal{F}_{\text{cam,R}}$. The color camera provides images with respect to $\mathcal{F}_{\text{cam,C}}$. Depth images are associated with the left camera frame $\mathcal{F}_{\text{cam,L}}$. The omnidirectional cameras are referenced to the frames $\mathcal{F}_{\text{omni,U}}$ and $\mathcal{F}_{\text{omni,D}}$ for the upward and downward facing cameras, respectively.

The frame definitions associated with the GNSS ground truth are illustrated in the two Figures 4 and 6a. The GNSS raw data describes the position of the left and right GNSS antenna frames $\mathcal{F}_{\text{L}}$ and $\mathcal{F}_{\text{R}}$ with respect to the world, that is, the topocentric frame $\mathcal{F}_{\text{T}}$. Note that GNSS data only provides positional information and no orientation, however we define the corresponding GNSS frames for the sake of completeness. Section 4.3 gives details regarding raw GNSS data processing to calculate the ground truth of the SUPER pose, that is, orientation and position of the central body frame $\mathcal{F}_{\text{B}}$ with respect to the world frame $\mathcal{F}_{\text{T}}$. This body frame is located above the IMU frame, precisely in the middle of the two GNSS antennas and is used as a central reference, also for the navigation algorithms in Section 6.

The topocentric reference frame $\mathcal{F}_{\text{T}}$ for each experiment is defined as the respective position of the GNSS base station. Finally, $\mathcal{F}_{\text{B,start}}$ defines the starting point of each trajectory and denotes the position of the body frame at time t_0, with t_0 being the starting time of the respective experiment. All relevant transformations between these reference frames are included in MADMAX for each experiment.

4 | EXPERIMENT SETUP

In this section, we discuss our field experiment methodology. The variability of conditions and the limited availability of equipment complicate the data set acquisition during field tests significantly, as opposed to well-defined laboratory environments. To structure the data acquisition process in such conditions and to facilitate follow-up data evaluation, we use predefined procedures for the experiment and the ground truth during all experiments.

4.1 | Sensor calibration

Similar to our previous data sets recorded on Mt. Etna in 2017 (Vayugundla et al., 2018), we performed intrinsic and extrinsic calibration of navigation cameras using the calibration toolbox DLR CalDe and DLR CalLab (see Strobl & Hirzinger, 2006 and Strobl et al., 2020, for the software) and IMU-to-cameras calibration as in Fleps et al. (2011). Due to the added omnidirectional stereo bench, the intrinsic and extrinsic calibration data sets are captured tying the configuration of all five cameras tightly into one joint calibration setup.

Intrinsic and extrinsic camera calibration is performed before the first experiments. We record images of the DLR CalDe calibration pattern (Strobl & Hirzinger, 2011) using the stereo cameras, the color camera, and two omnidirectional cameras. The derived camera

TABLE 2 List of frames used on SUPER

Frame	Description	X-axis	Y-axis	Z-axis
$\mathcal{F}_{\text{T}}$	Topocentric frame	East	North	Up
$\mathcal{F}_{\text{B}}$	SUPER body frame	Forward	Left	Up
$\mathcal{F}_{\text{B,start}}$	body frame at t_0	Forward	Left	Up
$\mathcal{F}_{\text{L}}$	Left GNSS antenna	Forward	Left	Up
$\mathcal{F}_{\text{R}}$	Right GNSS antenna	Forward	Left	Up
$\mathcal{F}_{\text{IMU}}$	IMU frame	Forward	Left	Up
$\mathcal{F}_{\text{cam,L}}$	Left stereo camera	Image horizontal	Image vertical	Optical axis
$\mathcal{F}_{\text{cam,R}}$	Right stereo camera	Image horizontal	Image vertical	Optical axis
$\mathcal{F}_{\text{cam,C}}$	Color camera	Image horizontal	Image vertical	Optical axis
$\mathcal{F}_{\text{omni,U}}$	Omnicam up	Image horizontal	Image vertical	Optical axis
$\mathcal{F}_{\text{omni,D}}$	Omnicam down	Image horizontal	Image vertical	Optical axis

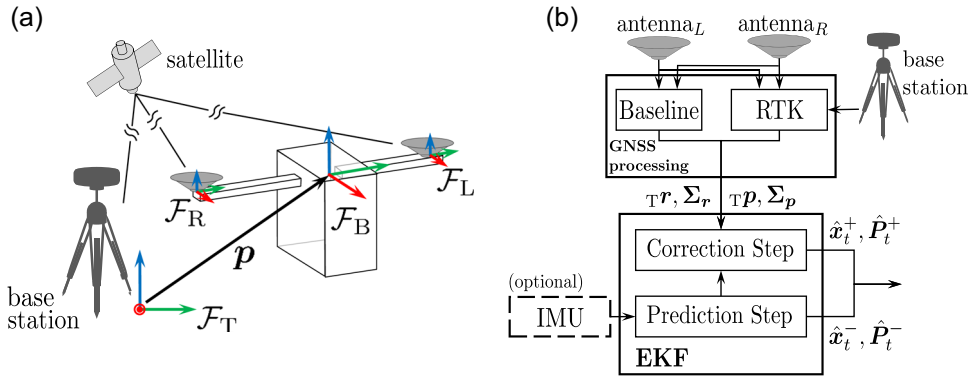


FIGURE 6 Illustration of the SUPER ground truth computation set-up and the processing steps for the GNSS-based precise positioning and attitude determination. The ground truth computation allows for optional use of the IMU. (a) Geometrical setup with relevant reference frames; and (b) processing pipeline. EKF, extended Kalman filtering; GNSS, global navigation satellite system; IMU, inertial measuring unit; RTK, real-time kinematic [Color figure can be viewed at wileyonlinelibrary.com]

parameters of the stereo system are consequently used for image rectification and depth map computation during all Morocco experiments. The raw images of the calibration are included in the data set, together with the derived intrinsic and extrinsic camera parameters for all cameras.

4.2 | Experiment procedures

We apply predefined procedures to each experiment to ensure consistency. We use a static platform as a base for SUPER, where we start and finish each trajectory to ensure that SUPER is placed in the same position and orientation both times. This allows for well-defined trajectory evaluation criteria and could also be used as loop closures for SLAM.

The platform is leveled horizontally and oriented to the east using a spirit level and a compass. This procedure provides a rough initial alignment of the navigation results with the GNSS ground truth and facilitates later processing of the data. The initial seconds (between 8 and 42 s) of each data acquisition run are recorded in a stand-still configuration to obtain static sensor readings for sensor bias evaluation.

At each location, neither the platform nor the GNSS base station move. Therefore all runs from that location have common start and end points in image and ground truth data. This allows for navigation and mapping overlaps between the different trajectories.

4.3 | RTK GNSS ground truth

One crucial aspect for field tests is the ground truth. In laboratory setups, the ground truth is usually obtained by high-precision optical tracking systems. For outdoor scenarios, such tracking systems are rarely available and, instead, RTK GNSS constitutes the main information source. Our GNSS setup is shown in Figure 6a together with the corresponding reference frames. We use the GNSS information of both antennas to compute the pose of the body frame $\mathcal{F}_B$ of SUPER as outlined in this section. The data set contains the raw GNSS observations from the left, right, and base antennas as well as the computed ground truth solution. Two distinct ground reference solutions are made readily available:

- an inertial-independent (GNSS-only) 5 DoF solution, sampled every second,
- a GNSS+inertial 6 DoF reference, sampled at 100 Hz. It is obtained by fusing the IMU and GNSS measurements.

We recommend the use of the 5 DoF ground truth solution for an inertial-independent evaluation of visual-inertial algorithms. The 6 DoF ground truth shall be used in the remaining cases, especially for the evaluation of purely vision-based algorithms.

The ground truth estimation solves the navigation problem, for which the position, velocity and attitude of a moving (rigid) body are determined. The kinematic quantities relate two coordinate systems: (i) the frame whose motion is described, *body frame* $\mathcal{F}_B$; (ii) the frame with which that motion is respect to, denoted as *topocentric frame* $\mathcal{F}_T$. This study adopts the conventions recommended for rotation and reference frames from Barfoot (2017). Figure 6a provides an illustration of the aforementioned navigation frames.

The state estimate is expressed as a discrete-time state-space model. Thus, at the time t, the state is described by

$$\boldsymbol{x}_t = \left[{}_T\boldsymbol{p}_t^T, {}_T\boldsymbol{v}_t^T, \boldsymbol{q}_t^T \right]^T, \quad {}_T\boldsymbol{p}, \quad {}_T\boldsymbol{v} \in \mathbb{R}^3, \quad \boldsymbol{q} \in \mathcal{S}^3 \tag{1}$$

where the position ${}_T\boldsymbol{p}$ and velocity ${}_T\boldsymbol{v}$ are expressed in the topocentric frame and the quaternion $\boldsymbol{q} \coloneqq \cos(\theta/2) + \boldsymbol{u}\sin(\theta/2)$, where $\boldsymbol{u} = u_x i + u_y j + u_z k$ is an unitary axis and θ is a rotation angle, follows the Hamilton convention and expresses the body-to-topocentric rotation. The estimated state is denoted as $\hat{\boldsymbol{x}}_t$ and the corresponding covariance matrix as $\hat{\boldsymbol{P}}_t$.

The navigation problem is addressed based on a standard multiplicative (also known as error-state) extended Kalman filtering (MEKF) formulation (Markley, 2003), with Figure 6b depicting the

estimation workflow. The prediction step is based on a constant-velocity, constant-attitude model for the GNSS-only solution, while classical inertial integration is employed as prediction for the GNSS +IMU solution.

During the GNSS processing stage, the raw code and phase observations from left and right antennas, along with the base station are fed to the open source RTKlib (Takasu & Yasuda, 2009). First, the *RTK* module performs a least-squares (LS) adjustment for the position solution ${}_{\mathrm{T}}\boldsymbol{p}$ and its covariance matrix $\sum_{\boldsymbol{p}}$ (Teunissen & Montenbruck, 2017). The *Baseline* module estimates the inter-antenna baseline vector ${}_{\mathrm{T}}\boldsymbol{r}$ and its covariance matrix $\sum_{\boldsymbol{r}}$. The *SD* for horizontal and vertical positioning of such a setup is approximately 2 and 5 cm, respectively, for fixed solutions (Medina et al., 2021). A fixed solution denotes a position/baseline result for which the estimated integer carrier phase ambiguities are considered valid, which is generally the case for MADMAX.

The resulting GNSS-based attitude precision coarsely relates to the accuracy of the estimated baseline (i.e., the positioning error for each antenna) compared to the actual antenna baseline (Giorgi et al., 2012). The SUPER configuration has a baseline length of 1.28 m. Compared to the estimated baseline length, the attitude precision results mostly below 0.5° for roll and yaw estimates for our data set. Generally, the GNSS accuracy is denoted in the corresponding *SD*s in the ground truth data.

Since only two GNSS antennas were installed, attitude determination for the GNSS-only ground truth becomes an ill-posed problem. As a result, roll and yaw can be accurately estimated, while pitch is not observable, thus providing a ground truth in 5 DoF. We provide the MATLAB code for our two ground truth estimation approaches together with the data set online.

5 | DATA SET

The focus of our Morocco experiments was to gather relevant planetary-analog data for navigation and mapping that offers a variety, both in the type of the trajectory as well as in the type of the terrain. An overview of the locations (labeled A–H) is given in Table 3, alongside with a brief description of the terrain, images of the scene, and the corresponding plots of the GNSS ground truth trajectories.

5.1 | Planetary analog location description

MADMAX contains 36 trajectories recorded during eight data acquisition sessions, each in a separate location, in three general areas: Gara Medouar, also known as *Rissani 1* (locations A–C, see Preston et al., 2012, p. 73, for details), *Kess Kess* (locations D and E, see Preston et al., 2012, p. 74) east of the city of Erfoud, and *Maadid* (locations F–H) north of Erfoud. All locations are marked in the map of Figure 2. Table 3 features impressions from SUPER's color camera that show the geological character of each location.

The Rissani area provides rich geological features for navigation cameras—rocks, pebbles, and sand patches. The area is featured in the background of Figure 1. Several landmarks are visible in omnidirectional cameras, most dominantly the tabletop mountain. The terrain is mostly traversable for a typical planetary rover with occasional challenges. The area is generally barren with vegetation visible sparingly. Additionally, it is the main site of testing for Brinkmann et al. (2019) and one of the three acquisition sites of Lacroix et al. (2019).

The Kess Kess area consists of a ridge formed by carbonate mounds (Preston et al., 2012, p. 74) with a wide flat terrain in its vicinity. This site has previously featured a 4-week long Mars-analog mission campaign, where many scientific experiments were conducted under simulated Martian surface conditions (Groemer et al., 2014).

Our data from that site is the result of two distinct acquisition sessions. The first one at the location (D) was realized in front of the ridge. The area is flat, barren, and has many pebbles providing features of high saliency as seen in the top left image of Figure 3. For omnidirectional cameras, the nearby mound provides a plethora of landmark features.

Our second Kess Kess data acquisition location (E) is situated between the mounds of the ridge—see the top right image of Figure 3. Here, the geological situation is different as a variety of rocks and stones was present and the landmark richness for omnidirectional cameras was high. In terms of terrain steepness and the number of rocks, we consider the location to be virtually non-traversable for planetary rovers of current generation, thus we exploit the high mobility of SUPER and its carrier to obtain the navigation data there.

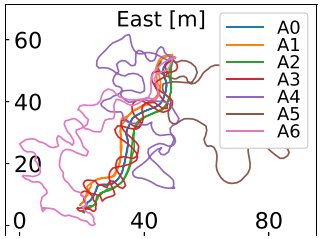
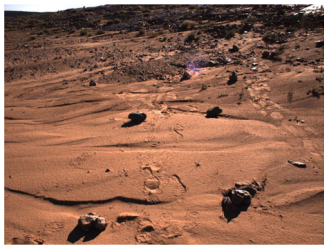
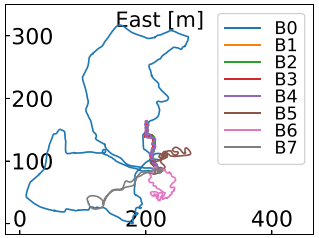
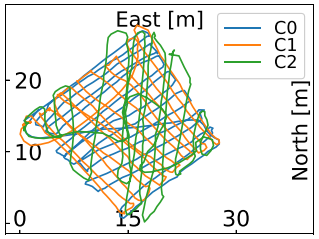
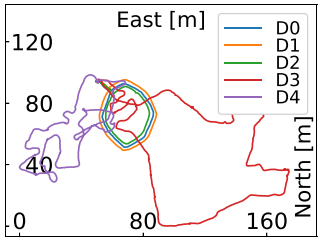
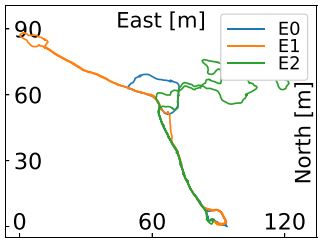
Lastly, the Maadid area hosts three distinct data acquisition sessions in very different terrain. The first one (F) is captured in a rather flat location covered with pebbles, easy to traverse for a planetary rover, as seen in the bottom left image of Figure 3. The environment for the second session (G) consists of a mixture of composite rock formations embedded in a sandy area. It comes with many landmarks for omnidirectional cameras to work with, and a terrain difficult to traverse for a planetary rover. The third data acquisition session (H) is situated in an area mostly consisting of sand dunes—see the bottom right image of Figure 3. These provide low saliency of corner or line features, thus being challenging for most visual odometries.

5.2 | Experiment properties

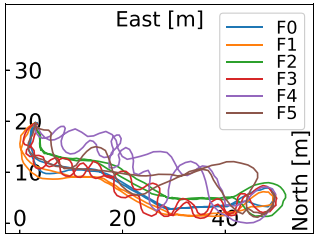
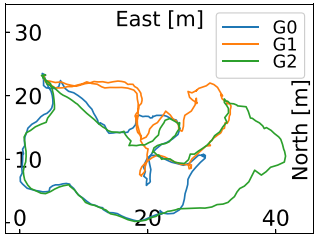
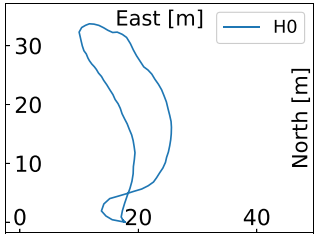
We made sure that all trajectories follow the same recording procedures, to make the data set more consistent. As outlined in Section 4.2, all tracks are realized in such a ways that the initial and final pose are the same—with heading always to the east, and roll and pitch approximately zero—allowing for at least one loop closure within each.

All runs which are obtained in one location have mutual identical starting and end poses, allowing for at least two overlaps between each pair of tracks for multirun mapping. Furthermore, all

TABLE 3 Experiment overview: GNSS trajectories, color camera impression of the scenes, and additional information sorted by location. Camera decalibration is indicated with

Trajectories	Scene	Character	ID	Length [m]	Type
		Flat area with stones, rock ridge at the end of the area. Low sun illumination.	A0	133	homing
			A1	138	homing
			A2	134	homing
			A3	219	zig-zag + homing
			A4	250	mapping
			A5	200	mapping
			A6	258	exploration
		Flat area with sand and pebbles and few big rocks, cliffs visible in the background.	B0	1511	long range nav
			B1	193	homing
			B2	195	homing
			B3	195	homing
			B4	286	zig-zag + homing
			B5	301	mapping
			B6	293	exploration
			B7	312	exploration
		Small flat and square area, half sandy and half stony.	C0	341	zig-zag
			C1	321	zig-zag
			C2	378	zig-zag
		Flat area with small stones and pebbles, hill formation in the background	D0	141	circular homing
			D1	155	circular homing
			D2	134	circular homing
			D3	493	long range nav
			D4	422	exploration run
		Rough terrain inside a valley and big stones within the traversed path.	E0	223	exploration
			E1	309	exploration
			E2	374	exploration
		Flat area with small pebbles, rough terrain at the end of the area.	F0	121	homing
			F1	128	homing
			F2	121	homing

(Continues)

Trajectories	Scene	Character	ID	Length [m]	Type
			F3 F4 F5	172 167 141	zig-zag + homing mapping mapping
		Navigation around big composite rock boulders with sandy surface in between.	G0* G1* G2*	125 115 154	exploration exploration exploration
		Desert sand dunes.	H0	90	exploration

trajectories are recorded in such way that they overlap on several additional occasions to allow for combined mapping.

For each experiment, we choose from predefined categories of trajectories that represent different aspects of navigation. Trajectories of the *mapping* type aim to cover an area with many overlaps within one run, trying to allow for dense terrain mapping. *Zig-zag* trajectories also aim at dense mapping of an area. In this case, a structured grid pattern of motion is used, unlike the unstructured motions of the mapping trajectories. *Long-range-navigation* runs cover long distances and are targeted towards evaluating localization algorithms. We also record trajectories for *homing* algorithms that follow one path several times with a minor offset. Finally, *exploration* runs combine several of the characteristics of the aforementioned types in an unstructured manner. The different trajectory types at each location are listed in Table 3, where the characteristics and mapping overlap of each run can also be seen in the corresponding GNSS ground truth overview plots.

The operator was instructed to move at a velocity, which is similar the movement speed of current or future planetary rovers and to keep this velocity constant. In the data set, the overall average velocity is at 29 cm/s, while the average velocities of the individual sequences range between 22 and 48 cm/s. F-0, with an average velocity of 12 cm/s, is considered as special case.

5.3 | Data set content

We provide sensor data from all sensors listed in Table 1. From the stereo cameras, we provide the rectified images of the left and right camera together with the resulting depth image. The color camera image stream includes rectified images as well. From the omnidirectional cameras, we provide the raw images. As the upper and bottom omnidirectional cameras show the face and the legs of the human carrier, an area of 60° is blacked out in each image as shown in Figure 5.

All images are named after the respective camera plus the timestamp of the UNIX time in nanoseconds, and indicate if the images are rectified, resulting in the following pattern: `img_{rect}_{camera}_{timestamp}.png`. We use the common Portable Network Graphics data type for the images. The time-stamped IMU data are provided as comma-separated values (`.csv`) files for each experiment. The stereo cameras are synchronized with the IMU, with the IMU being used to trigger the cameras. The two omnidirectional cameras provide synchronized image pairs. The other sensors run independently, however, all of SUPER's sensors use clock synchronization to provide precise timestamps in the system time.

The raw GNSS data from the left, right, and base antennas are provided as `.obs` RINEX format (International GNSS Service, 2015) together with the satellite ephemeris (`.nav`) text files. The post-processed ground truth pose of the SUPER body center from Section 4.3 is provided in an additional `.csv` text file as position vector, rotation in quaternions, a linear velocity vector, together with the associated *SD*s and the timestamp in nanoseconds.

The ground truth is provided in two ways, as the GNSS-only 5 DoF ground truth, and in addition, as the 6 DoF fusion of GNSS and IMU data. Note that the timestamps of the postprocessed GNSS measurements are temporally synchronized with the other sensor data of SUPER. The raw GNSS data is provided in UTC time according to the RINEX specification, thus the temporal synchronization with the other sensors has to be taken into account. The temporal offset between GNSS time and the UNIX system time is listed in the `metadata.yaml` file for each experiment. Note that this time offset is different for each day, as SUPER was not connected to clock synchronization servers during the field campaign.

Additionally, the metadata text file lists detailed information for each experiment, like precise location coordinates of the base station, the time stamps of the experiment start, and the start of data recording, respectively. One key information is the initial pose of SUPER with respect to the base station, that is, the transformation from $\mathcal{F}_{\mathrm{B,start}}$ to $\mathcal{F}_{\mathrm{T}}$.

In addition to the experiment data, we provide calibration data. This includes the intrinsic and extrinsic camera calibration as `callab_camera_calibration*.txt` together with the resulting camera parameters for the rectified images as `{camera}_rect_info.txt`. The transformations between the relevant coordinate frames from Table 2 are provided as well. This is a collection of transforms between a parent and a child frame, given as position and quaternion-orientation in `.csv` files.

The navigation results for ORB-SLAM2 and VINS-Mono from Section 6 are included as well. They are provided in two formats: The original result data with respect to the camera and IMU frames, respectively, and the data aligned with the GNSS ground truth. Both are text files with timestamp, position, and orientation quaternion for each pose.

Finally, we provide the data set from the SUPER calibration. This includes raw images from each camera with the calibration pattern visible, the specification of the calibration pattern dimensions, as well as the calibration results in a text file.

All data can be freely accessed online at https://rmc.dlr.de/morocco2018. The website shows details on each experiment location and the experiments performed, plus one section for the calibration data. The data is available individually for each experiment, structured as shown in Figure 7 and is provided in a compressed format.

5.4 | Issues, challenges, and lessons learned

Operations in extreme environments pose special challenges to the system, the operators, and the experiments. In our case, several challenges and technical issues were encountered, which we could partially account for.

Many recordings contain optical disturbances that make the data set challenging. One disturbance appeared in particular during afternoon experiments: lens flares due to a low sun position. Another disturbance was the strong over or underexposure of image regions

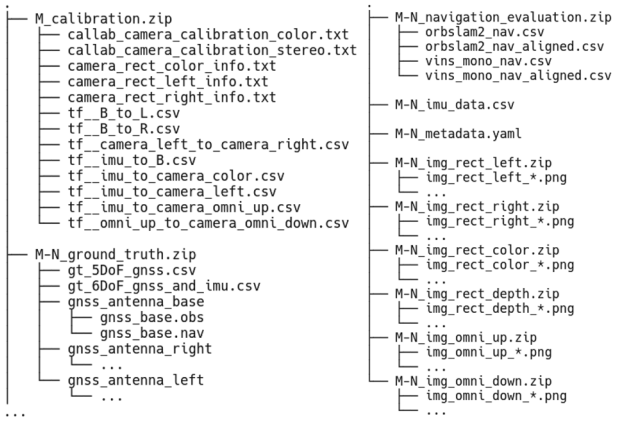


FIGURE 7 Example structure of the available data for an experiment from *Location M* with *Run ID N*. The data are provided in a compressed format

due to shadows in the field of view. The moving shadow of SUPER introduces the similar optical features in the recorded images like a planetary rover. Our analysis concluded that the human carrier did not introduce any additional undesired shadows into the image that might disturb navigation algorithms. All of these disturbances were desired to create challenging scenarios for planetary navigation algorithms as a robustness test. Due to the operator having moved slowly and exposure times in bright sunlight having been short, no significant motion blur was observed.

The first issue to mention is the extrinsic calibration of cameras. Once the field tests had been completed, we evaluated the quality of the extrinsic stereo camera calibration by comparing the vertical displacement of sampled features within selected stereo pairs of each run. It turned out that the last runs of the Morocco campaign, labeled G runs, experienced a vertical feature offset bigger than 2 px, which we consider a sign of decalibration. As a result, the computed depth images for these runs are less accurate and contain several invalid regions. Furthermore, the accuracy of camera to IMU calibration is degraded in these runs. The cause for this calibration error is most likely an unexpected mechanical load during transport to the experiment site. Nevertheless, we publish the G runs so that the data can be used to test the robustness of algorithms against extrinsic decalibration. Indeed, the Section 6 shows that VINS-Mono and ORB-SLAM2 obtain accurate navigation results for the G runs. All other runs turned out to have accurate calibration. For future field tests, we recommend calibrating cameras, and IMU-to-camera, on a daily basis to ensure high data quality.

Throughout the field test, we experienced network problems that specifically affected the stereo cameras connected via gigabit Ethernet (GigE vision®). As a result, several frame drops occurred. These frame drops usually lasted for one to four consecutive frames (up to a quarter of a second) and seldom reached half a second, that is, up to eight consecutive missing frames. Our analysis shows that this still accounts for an inter-frame overlap of 80%–90% and 70%–80%, respectively.

Most runs experience frame drops of only 5%–10% of the overall frame count. However, the F runs are strongly affected with a loss of 15%–19%. We attribute the issue to the network hardware used in SUPER, which was chosen due to its lightweight design. Reconfiguring the network settings made the issue less prominent, but the general problem still prevailed. Generally, no direct correlation was found between the number of frame drops and poor navigation results, which we discuss in Section 6 in detail. The individual losses per run are listed together with the data. To overcome frame drops in the future, a more robust network setting has to be considered, even though this would require more heavyweight components to be used.

Finally, our GNSS solution lost precision in its measurements occasionally (we consider position measurements with a *SD* of more than 0.06 m to be imprecise) for two reasons:

- The RTK GNSS quality depends significantly on having a precise geo-reference solution of the base station. During the G and H runs, we recorded the base station GNSS data for intervals that were too short to obtain a sufficiently precise GNSS base station solution, thus leading to poorer accuracy in the corresponding pose estimation of SUPER. For future experiments, a prolonged data recording for the GNSS base station should be considered in the experiment schedule.
- During some experiment runs, the SUPER antennas lost the signals of several GNSS satellites for a few seconds. On such occasions, the precision was usually good enough to provide a satisfactory position estimate, but the orientation suffered significantly.

The GNSS inaccuracies occurred in 11% of all measurement points, not counting the runs B7, C2, and F0, which were more strongly affected with rates exceeding 40%. The accuracy of the measurements is represented in the GNSS pose estimate *SD* for each timestamp. Any algorithm or any evaluation that also considers the associated uncertainties should not be affected by this issue.

6 | EVALUATION WITH STATE-OF-THE-ART NAVIGATION ALGORITHMS

We evaluated the data set using two state-of-the-art SLAM-based navigation algorithms. The algorithms provide us with a 6D-pose of the SUPER system, which is subsequently compared with the GNSS ground truth. The motivation is to provide the navigation results and insights as a baseline against which other algorithms can be compared, and we invite interested researchers to do so.

6.1 | VINS-mono and ORB-SLAM2 SLAM setup

Among the different motion estimation algorithms, visual odometry (VO) and SLAM are the processes of concurrently estimating the

map of the environment as well as the robot pose (position and orientation) from a stream of images. VO incrementally estimates the path of the camera/robot focusing mainly on local consistency, whereas SLAM obtains a globally consistent estimation of the camera/robot trajectory and map by recognizing previously mapped areas (loop closure). The extensive navigation sequences covered by MADMAX containing numerous loop closure opportunities are suitable for testing both VO and SLAM algorithms.

VO and SLAM can work with just a single monocular camera. The cameras allow robust and accurate place recognition, and thanks to their small size, low cost, and easy hardware setup they are of great interest in the robotics community. However, there is a list of drawbacks limiting the use of monocular cameras in real-world robotic applications. Monocular vision-only systems cannot recover the metric scale of the scene, they also suffer from scale drift, and pure rotation movements cause VO systems to fail during exploration. In addition, the initial map required for system bootstrapping cannot be obtained from the first frame, so multiview or additional sensors are required to produce it. A wide variety of state-of-the-art techniques is available to address these problems, especially in the field of sensor fusion. Two common VO/SLAM configurations that overcome the challenges while also taking advantage of the features offered by MADMAX are: stereo and visual inertial.

The scene structure can be reliably obtained with stereo cameras through static triangulation with depths within a range of ~40 times the stereo baseline (Mur-Artal & Tardós, 2017), that is, 3.6 m for our case. The possibility of having the structure of the scene instantly yields a true-scale SLAM solution, solving among other problems the procedure of initialization from unknown initial states.

A monocular VINS consists of a camera and an IMU. One advantage of this setup is to observe the metric scale, as well as directly measuring roll and pitch angles. To estimate a valid scale, the robot has to experience a nonconstant acceleration, which is the case for most moving robots. Furthermore, cameras allow for accurate assessment of slow movements, whereas IMUs are well suited for observing of fast movements and rotation. In practice, an IMU is a valuable complement to the visual data, since compared to cameras, these sensors are independent of the environment; their high-rate values are also cheap to process and have an accurate probabilistic model with little to no outliers.

SLAM approaches are used to counteract the long-term drift in the translation and orientation that can strongly affect the visual odometry navigation. SLAM systems can detect online when the rover returns to a mapped area (place recognition and loop closure modules) and correct the drift accumulated in the exploration (graph optimization and/or bundle adjustment). The relocalization of the camera after a tracking failure (due to lighting changes, aggressive movement or lack of a textured scene) produces a very robust and zero-drift tracking method. For all these reasons, the place recognition and pose-graph optimization are key modules that play an important role to operate in large environments as it is the case for MADMAX.

As we mention above, we test MADMAX with two different state-of-the-art SLAM baselines: ORB-SLAM2 (Mur-Artal & Tardós, 2017) and VINS-Mono (Qin et al., 2018). ORB-SLAM2 for stereo cameras is built on monocular feature-based ORB-SLAM (Mur-Artal et al., 2015), a complete system for monocular cameras, including map reuse, loop closing, and relocalization capabilities. VINS-Mono is a tightly coupled monocular visual inertial odometry that fuses IMU measurements and feature observations. Both systems work in real time on standard CPUs in a extensive variety of environments from small hand-held sequences, to ground robots and drones. An open-source system integration is available for both algorithms and their performances have been validated on public data sets and real-world experiments.

6.2 | Evaluation results

MADMAX is a large-scale data set that provides suitable sequences to test stability and long-term use of SLAM. Notice that results and evaluation shown in this section do not aim to compare performance specifically between ORB-SLAM2 and VINS-Mono but to evaluate general differences between stereo and visual inertial setups using the selected algorithms as respective examples. Their performance is also used to show the opportunities provided by the data set in the sense of navigation algorithm testing. Additionally, the evaluation aims to provide a navigation baseline for the respective category, which can be used to benchmark other algorithms. The SLAM algorithms compute 6D-poses for every frame of the sequences that we compare with the GNSS ground truth. Next, we will list the details of our evaluation:

- Both systems have been tested using half-resolution images from the monochrome cameras to achieve real-time performance using our institute computers (Intel Xeon E5-1630, 3.70 GHz, 16 GB RAM, CPU-only computation).
- We initialize VINS-Mono with an estimation of the extrinsic calibration parameters from the initial calibration and let the system refine them online.
- The association with the GNSS data has been performed by only considering GNSS measurement points with a *SD* lower than 0.06 m.
- For evaluation, we use the absolute trajectory error (ATE) and the relative pose error (RPE), as proposed by Sturm et al. (2012).
- We consider the fully optimized trajectories that use all data available at the end of each run.
- We use ORB-SLAM2 and VINS-Mono with loop closing and relocalization capabilities enabled for each individual sequence, but without map reuse between the runs.

The ATE calculates the root-mean square error (RMSE) of all global positions $\boldsymbol{p}_t$ along the frames of the estimated trajectory with respect to the GNSS ground truth correspondences $\tilde{\boldsymbol{p}}_t$, after both

curves have been aligned using the method from Horn (1987). The resulting error at timestep t is

$$\mathrm{ATE}_t = \| \boldsymbol{p}_t - \tilde{\boldsymbol{p}}_t \| \tag{2}$$

and the overall ATE is

$$\mathrm{ATE} = \sqrt{\sum_{t=0}^{N} (\boldsymbol{p}_t - \tilde{\boldsymbol{p}}_t)^2 / N} \tag{3}$$

for a total of N trajectory segments.

The RPE computes the RMSE of the difference of traveled distances between the estimated trajectory and the ground truth. The traveled distance between two frames separated temporally by Δt is defined as $d_t = \| \boldsymbol{p}_t - \boldsymbol{p}_{t+\Delta t} \|$ and the resulting RPE as

$$\mathrm{RPE} = \sqrt{\sum_{t=0}^{N} (d_t - \tilde{d}_t)^2 / N}, \tag{4}$$

where we choose $\Delta t = 1$ s. The step-wise RPE is therefore

$$\mathrm{RPE}_t = \| d_t - \tilde{d}_t \|. \tag{5}$$

Note that while on the one hand, ATE computes the absolute difference between the two trajectories in meters, RPE evaluates the average pose drift in meters per second.

Since, on occasion, one of the SLAM approaches may not be capable of calculating the complete trajectories, for a fair comparison, we use these metrics just when at least 75% of the trajectory traveled distance has been accomplished successfully. The exemplary results for four trajectories of SUPER computed by ORB-SLAM2 and VINS-Mono are illustrated in Figure 8 together with the GNSS ground truth.

6.2.1 | Navigation robustness

To evaluate the navigation robustness, the estimated percentage of accomplished trajectories is shown in Figure 9 for each sequence. It can be seen how the visual inertial navigation is more robust than stereo since it finishes most of the sequences, a significantly greater number than ORB-SLAM2. It turned out that the frame drops in the recordings, as mentioned in Section 5.4, do not have a direct correlation with the navigation robustness of the algorithms. This is shown, for example, by the run F2 (also shown in Figure 8), which has one of the highest frame drop rates at 19% but VINS-Mono and ORB-SLAM2 complete the full trajectory with a very low error. It is due to the fact that the inter-frame overlap of 70%–90% at such frame drop occurrences is sufficient for a continuous tracking for both algorithms.

On the other hand, both algorithms fail to navigate with A6, a run that has only a 5% rate of frame drops, as well as only a small number of consecutive dropped frames (3–4 dropped frames in a row). It turned out that *critical events*, that is, the combination of a low inter-frame overlap of 70%–90% (e.g., due to frame drops) together with other disturbances can cause tracking and relocation failures.

Such disturbances are the change of the exposure time, over and underexposure of parts of the image, lens flares, back light, or the

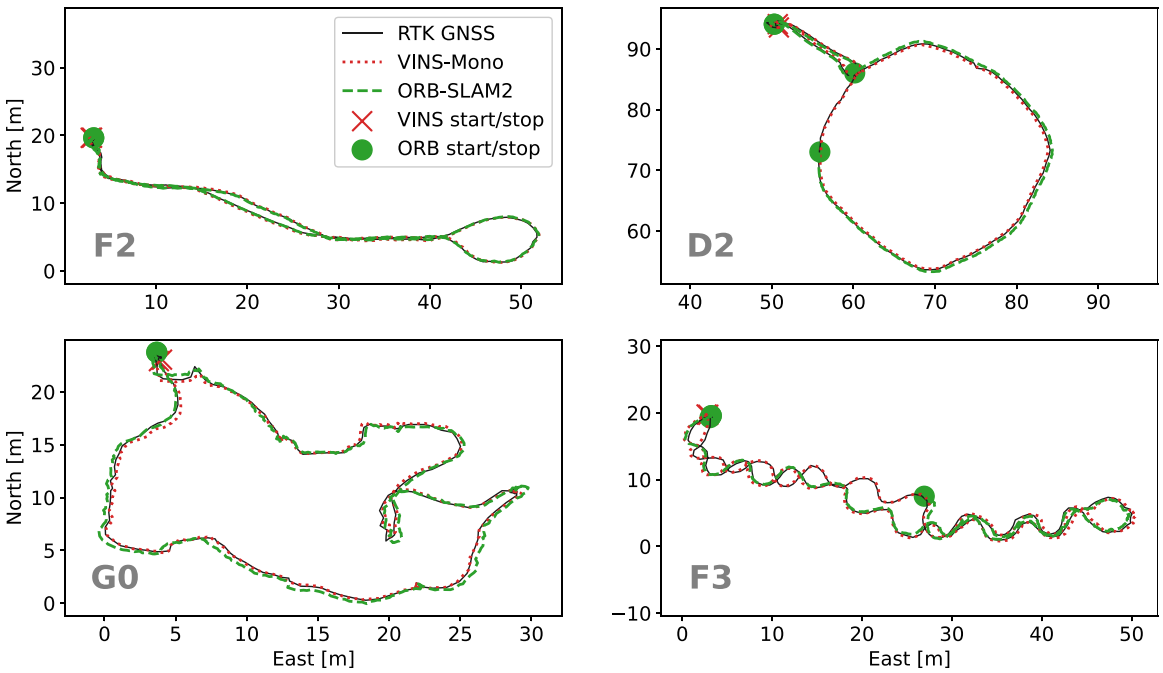


FIGURE 8 Sample results of the navigation: The performance of ORB-SLAM2 and VINS-Mono with respect to the GNSS is shown for four experiments. Ideal results (F2), ORB-SLAM2 loses track but relocates after recognizing a previously visited area (D2), successful navigation of both algorithms with the extrinsic decalibration (G0), and loss of ORB-SLAM2 tracking (F3) are shown. GNSS, global navigation satellite system [Color figure can be viewed at wileyonlinelibrary.com]

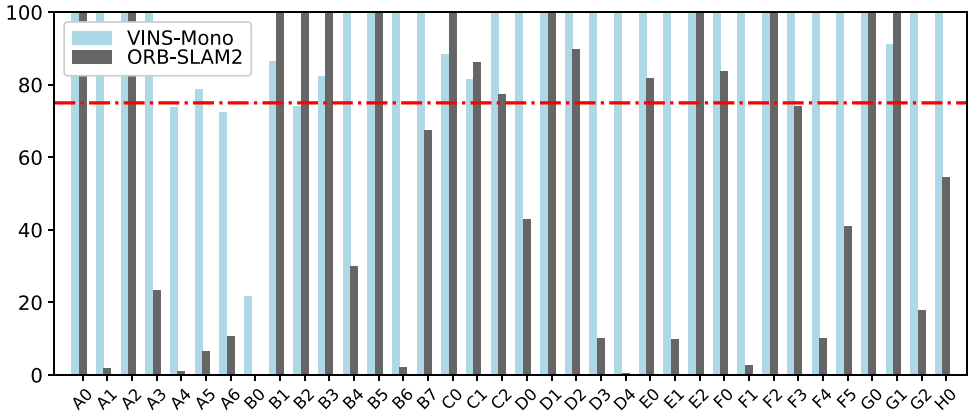


FIGURE 9 Percentage of each trajectory completed by ORB-SLAM2 and VINS-Mono. Only navigation results with more than 75% completion (red line) are considered in the error analysis [Color figure can be viewed at wileyonlinelibrary.com]

moving shadow of SUPER. Since both algorithms are feature-based approaches, they deal with purely visual noise in a similar way. In our case, it turned out that ORB-SLAM2 normally overcomes disturbances when the overlap between consecutive frames is greater than ~80%. Values below that threshold in combination with challenging visual conditions cause tracking and relocation failures. As might be expected, VINS-Mono is more robust against tracking failures produced by visual effects. Since the higher IMU measurement rate allows for a continuous pose update between two consecutive camera frames, VINS-Mono is able to bridge the gaps and therefore keep the functioning of the tracking thread intact.

In general, these critical events can occur at any time during a sequence, thus explaining the seemingly arbitrary difference in completed trajectory length for each experiment. Also note that the runs A3–A6 were recorded in the late afternoon (around 4 pm in December) with cameras facing the direction of the setting sun. Thus, being recorded under more challenging illumination conditions is reflected in the poor navigation outcomes in terms of the trajectory completion percentage. Sequences F3 and D2 in Figure 8 show the case of loss of tracking for ORB-SLAM2, where relocation succeeds for the D2 run and fails for the F3 run.

6.2.2 | Navigation accuracy

Leaving aside the fact that the visual inertial algorithm manages to complete more sequences than its stereo-based counterpart, it also performs slightly better in terms of ATE accuracy, as shown in Figure 10. Evaluating the 16 sequences in which both algorithms reach more than 75% completion, VINS-mono outperforms ORB-SLAM2 in 10 sequences, whereas ORB-SLAM2 performs better in six runs. Nevertheless, ATEs for both systems are within the same range.

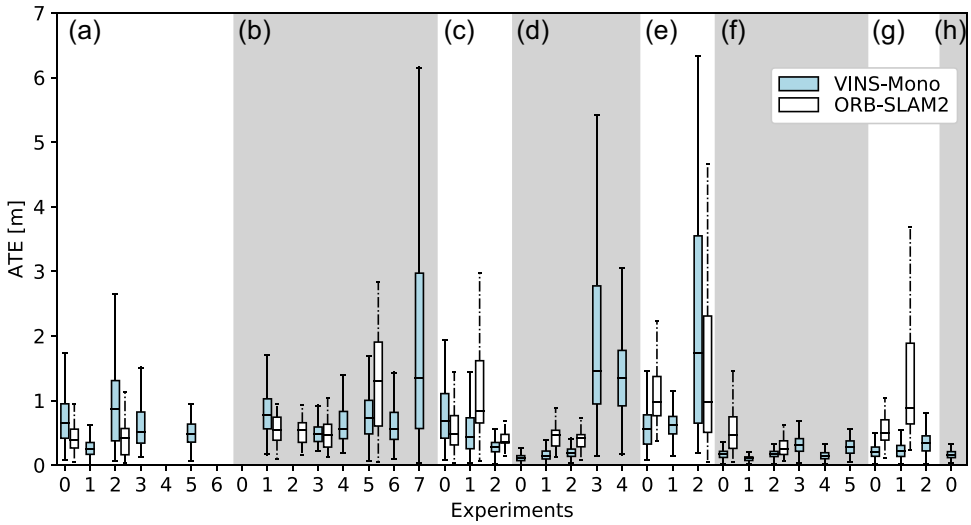


FIGURE 10 Distribution of the step-wise ATE of ORB-SLAM2 and VINS-Mono for each experiment. Only navigation results with a completion of the trajectory of at least 75% are considered. The whiskers on each vertical bar denote the minimum and maximum values of the error distribution for each run. The box denotes the first and third quartile of the data with the median as the dividing line in-between. ATE, absolute trajectory error [Color figure can be viewed at wileyonlinelibrary.com]

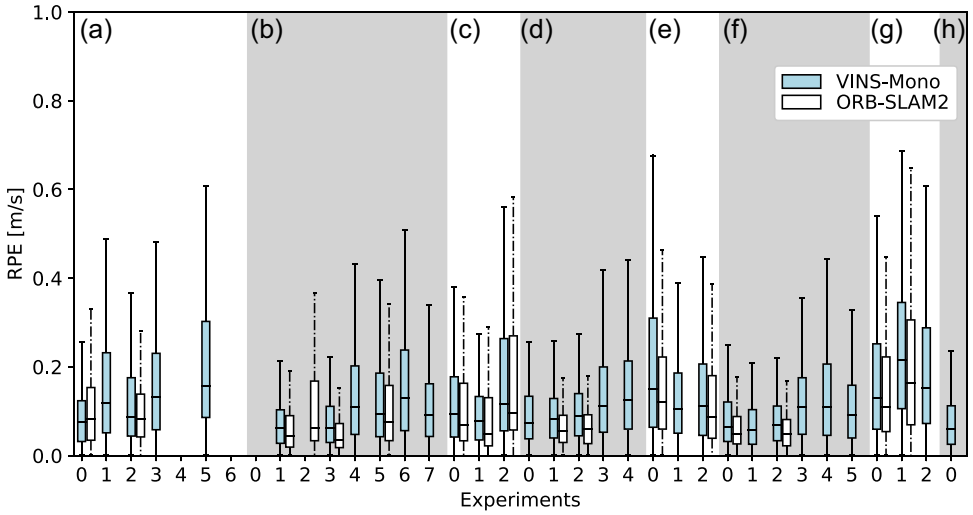


FIGURE 11 Distribution of the step-wise RPE of ORB-SLAM2 and VINS-Mono for each experiment. Only navigation results with a completion of the trajectory of at least 75% are considered. The whiskers on each vertical bar denote the minimum and maximum values of the error distribution for each run. The box denotes the first and third quartile of the data with the median as the dividing line in-between. RPE, relative pose error [Color figure can be viewed at wileyonlinelibrary.com]

On the other hand, ORB-SLAM2 outperforms VINS-mono in 15 out of the 16 sequences in terms of RPE as shown in Figure 11. Generally, it can be said that both algorithms provide accurate navigation results with minor differences with respect to ATE and RPE accuracy.

We compare the performance of ORB-SLAM2 and VINS-Mono in MADMAX with respect to the EuRoC data set (Burri et al., 2016). The EuRoC data set contains sequences recorded by a MAV flying around different indoor environments, and both ORB-SLAM2 and VINS-Mono authors provide testing values from the resulting ATE for their systems. To remove the effect of different trajectory lengths on the resulting alignment error between EuRoC and MADMAX, we compute the median of the ATEs normalized with the length of each sequence in Table 4. The accuracy of EuRoC and MADMAX in terms of ATE lies at similar levels.

Nevertheless, a degradation in performance occurs from EuRoC to MADMAX that cannot be attributed to longer trajectories. There are several potential explanations for this, for example, the challenging visual content of the sequences or the different quality of the sensor calibration.

Nevertheless, the resulting navigation results of both algorithms can be considered as accurate, see for example, the trajectories of run F2, which is shown in the Figure 8. Until tracking is lost, ORB-SLAM2 also provides accurate navigation results for the sequences F3 and D2 (Figure 8). The same figure also shows the G0 run, which is one of the three runs with extrinsic decalibration. It is clear that both algorithms cope with such decalibration and provide reliable navigation results. Apart from the obvious advantages of visual inertial SLAM versus stereo SLAM in terms of robustness for outdoor environments with long-term trajectories, we have not been able to observe any major differences between the two state-of-the-art SLAM pipelines.

TABLE 4 Normalized absolute trajectory error (%)

	VINS-Mono	ORB-SLAM2	~ Length (m)
EuRoC	0.16	0.27	80
MADMAX	0.27	0.35	175

Note: Comparison of the navigation performance of the EuRoC data set and MADMAX.

6.3 | Comparison of hand-held and rover-based navigation

Finally, we investigate how representative the hand-held data is for planetary rover navigation, answering the question of whether human-induced motions negatively affect the navigation algorithms. We take seven navigation sequences that were obtained by the rover-mounted SUPER unit and test these using ORB-SLAM2 and VINS-mono. We apply identical evaluation methods to those in Section 6.2. Note that these data belong to the *InFuse* project (Post et al., 2018) and are therefore not included in MADMAX.

The seven sequences are 26–54 m in length. The trajectories are mostly straight drives combined with wide curves and took place in, and close to, the area of the C runs. Therefore, both MADMAX and these seven rover based sequences experience close to identical environmental conditions. The average velocity of the rover-mounted SUPER is 12 cm/s, about a third of the velocity of the hand-held navigation (at 29 cm/s). The position of the stereo cameras is approximately 0.80 m above the ground.

The main difference in both data sets is therefore the different type of movement, where the human-induced motions may influence the navigation in a negative way. Recall that we experienced no

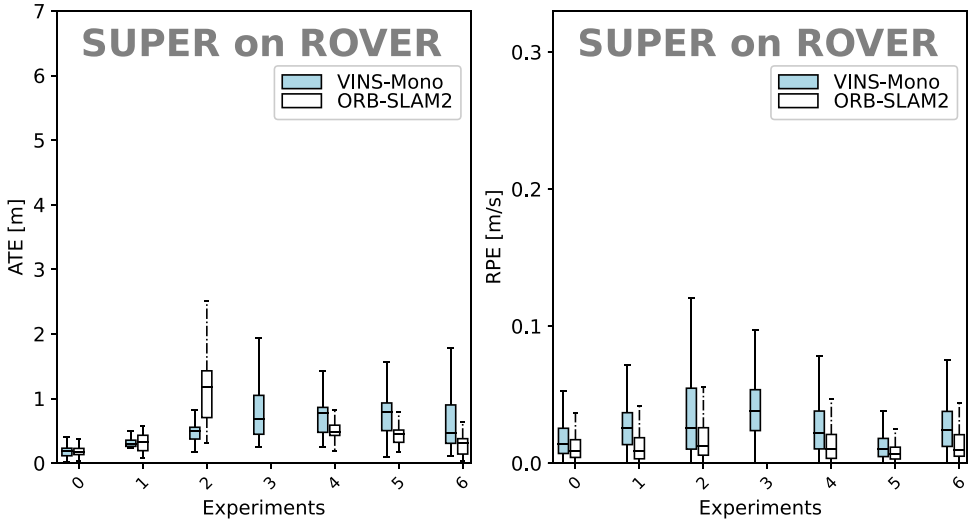


FIGURE 12 Distribution of ATE and RPE of ORB-SLAM2 and VINS-Mono for the rover-mounted SUPER sequences. Note that the RPE is scaled by a factor of 3 compared to the previous figure to account for the lower velocity. The whiskers on each vertical bar denote the minimum and maximum values of the error distribution for each run. The box denotes the first and third quartile of the data with the median as the dividing line in-between. ATE, absolute trajectory error; RPE, relative pose error [Color figure can be viewed at wileyonlinelibrary.com]

motion blur in neither of the two data sets, thus limiting the difference solely to the type of motion.

We make the claim that MADMAX can be considered a representative data set for rover navigation if the ATE and RPE of the experiments match the errors of the rover-mounted SUPER experiments. We consider the C runs and the D-0 to D-2 runs for a comparison, as these were obtained in the same location or feature similar types of trajectories, respectively. Recall, that the RPE depends on the experiment velocity, as it is the distance error per time as stated in (5). We therefore expect it to be lower by a factor of three for the rover-bound experiments.

Regarding the rover-based navigation, ORB-SLAM2 completes six sequences, except for run 3, whereas VINS-mono completes all seven runs. Figure 12 shows the respectiv navigation results in terms of ATE and RPE. Indeed, the ATE lies in the same range as the comparable runs from MADMAX, generally around 0.5 m with peaks at 2–3 m. The RPE is approximately one-third compared to the hand-held runs, which is expected owing to the three-fold difference in velocity. We therefore conclude that motions from the human-based transportation do not negatively influence the navigation. This indicates that MADMAX consists of representative planetary rover navigation data, supporting our case in favor of hand-held field testing.

7 | CONCLUSION

In this paper, we presented a field testing approach for planetary robotics navigation and mapping algorithms and test data recording that fills the gap between laboratory tests and complex, full-rover-system field tests. To do this, we deployed a compact hand-held sensoric abstraction of a planetary rover—SUPER—in a Mars-analog environment in the northern Sahara in Morocco. The result of the field test is the comprehensive Mars-analog VINS data set MADMAX that we make publicly available.

This data set includes recordings of monochrome stereo cameras, a color camera, two omnidirectional cameras in a vertical stereo setup, and an IMU. The experiments took place in several distinctive locations, and we outlined the variety and character of the different experiments. We discussed several operational aspects that turned out to be crucial for a successful data set recording, such as the ground truth computation of position and orientation from the GNSS data, procedures for data recording, and the calibration of five different cameras relative to each other, including the two omnidirectional cameras.

Finally, we showed that the recorded data can be used for navigation by applying the state-of-the-art algorithms ORB-SLAM2 and VINS-Mono. We evaluated their performance for this planetary-analog setting, showed their mostly high accuracy, but also revealed corner cases were these algorithms fail. We compare the performance of the algorithms to a rover-based data set and show that our hand-held approach does not negatively influence the accuracy of the state of the art.

It became apparent that MADMAX is a challenging data set for planetary navigation which can be used as robustness test and performance reference for new navigation approaches. We make the data publicly available and provide detailed information about it to facilitate the use of the recordings.

ACKNOWLEDGMENTS

We want to thank all European Space Robotics Technologies Research Cluster participants in the Morocco 2018 field test for their help in the field, their valuable scientific input, and their general support. This study was funded by the DLR project *MOdulares Robotisches EXplorationssystem* (MOREX). This activity has been conducted jointly with the two European Commission Horizon 2020 Projects InFuse and Facilitators.

They received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 730068 and 730014. We want to thank our colleague Klaus Strobl for his valuable ideas and feedback on the paper and our colleagues Bertram Willberg, Tim Bodenmüller, Bernhard Vodermayer, and many more for their help in preparing the field campaign. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

All data that is considered part of the publication is publicly available under the URL https://rmc.dlr.de/morocco2018. A registration form needs to be filled out before accessing the data. An exception to this is the evaluation in Section 6.3 of the paper, as these data belong to a consortium of project partners.

ORCID

Lukas Meyer http://orcid.org/0000-0001-9514-8494
Laura Oliva Maza https://orcid.org/0000-0001-5382-3025
Daniel Medina https://orcid.org/0000-0002-1586-3269
Martin J. Schuster https://orcid.org/0000-0002-6983-3719
Mallikarjuna Vayugundla https://orcid.org/0000-0002-9277-0461
Marcus G. Müller https://orcid.org/0000-0003-4283-6693
Armin Wedler https://orcid.org/0000-0001-8641-0163
Rudolph Triebel https://orcid.org/0000-0002-7975-036X

REFERENCES

Antonini, A., Guerra, W., Murali, V., Sayre-McCord, T., & Karaman, S. (2020). The Blackbird UAV dataset. *The International Journal of Robotics Research*, *39*(10–11), 1346–1364.

Bakambu, J., Langley, C., Pushpanathan, G., MacLean, W., Mukherji, R., & Dupuis, E. (2012). Field trial results of planetary rover visual motion estimation in Mars analogue terrain. *Journal of Field Robotics*, *29*(3), 413–425.

Barfoot, T. D. (2017). *State estimation for robotics*. Cambridge University Press.

Brinkmann, W., Cordes, F., Koch, C. E. S., Wirkus, M., Dominguez, R., Dettmann, A., Vögele, T., & Kirchner, F. (2019). Space Robotic Systems and Artificial Intelligence in the Context of the European Space Technology Roadmap. *Proceedings of Space Tech Conferences*.

Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W., & Siegwart, R. (2016). The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*, *35*(10), 1157–1163.

Bussmann, K., Meyer, L., Steidle, F., & Wedler, A. (2018). Slip modeling and estimation for a planetary exploration rover: Experimental results from Mt. Etna. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2449–2456).

Cordes, F., Kirchner, F., & Babu, A. (2018). Design and field testing of a rover with an actively articulated suspension system in a Mars analog terrain. *Journal of Field Robotics*, *35*(7), 1149–1181.

Corke, P., Strelow, D., & Singh, S. (2004). Omnidirectional Visual Odometry for a Planetary Rover. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4007–4012). IEEE.

Fleps, M., Mair, E., Ruepp, O., Suppa, M., & Burschka, D. (2011). Optimization based IMU camera calibration. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3297–3304). IEEE.

Furgale, P., Carle, P., Enright, J., & Barfoot, T. (2012). The Devon Island rover navigation dataset. *The International Journal of Robotics Research*, *31*(6), 707–713.

Geromichalos, D., Azkarate, M., Tsardoulias, E., Gerdes, L., Petrou, L., & Perez Del Pulgar, C. (2020). Slam for autonomous planetary rovers with global localization. *Journal of Field Robotics*, *37*(5), 830–847.

Giorgi, G., Teunissen, P. J., & Gourlay, T. P. (2012). Instantaneous global navigation satellite system (GNSS)-based attitude determination for maritime applications. *IEEE Journal of oceanic engineering*, *37*(3), 348–362.

Groemer, G., Soucek, A., Frischauf, N., Luger, M., Stumptner, W., Ragonig, C., Sams, S., Bartenstein, T., Häuplik-Meusburger, S., Petrova, P., Evetts, S., & Sivenesan, C. (2014). The MARS2013 Mars analog mission. *Astrobiology*, *14*(5), 360–376.

Gunes-Lasnet, S., Winnendael, M., Diaz, G., Schwenzer, S., Pullan, D., Kisdi, A., Josset, J.-L., Ciarletti, V., Barnes, D., Griffiths, A., Paar, G., Allouis, E., Waugh, L., Woods, M., & Shaw, A. (2014). SAFER: The promising results of the Mars mission simulation in Atacama, Chile. *Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space*. ESA.

Hewitt, R. A., Boukas, E., Azkarate, M., Pagnamenta, M., Marshall, J. A., Gasteratos, A., & Visentin, G. (2018). The Katwijk beach planetary rover dataset. *The International Journal of Robotics Research*, *37*(1), 3–12.

Hidalgo-Carrió, J., Poulakis, P., & Kirchner, F. (2018). Adaptive localization and mapping with application to planetary rovers. *Journal of Field Robotics*, *35*(6), 961–987.

Hirschmüller, H. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(2), 328–341.

Horn, B. K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, *4*(4), 629–642.

International GNSS Service. (2015). RINEX: The receiver independent exchange format. https://igs.org/pub/data/format/rinex303.pdf

Lacroix, S., De Maio, A., Labourey, Q., PaivaMendes, E., Narvor, P., Bissonette, V., Bazerque, C., Souvannavong, F., Viards, R., & Azkarate, M. (2019). The Erfoud dataset: A comprehensive multi-camera and Lidar data collection for planetary exploration. *Proceedings of the Symposium on Advanced Space Technologies in Robotics and Automation*.

Lamarre, O., Limoyo, O., Marić, F., & Kelly, J. (2020). The Canadian planetary emulation terrain energy-aware rover navigation dataset. *The International Journal of Robotics Research*, *39*(6), 641–650.

Lutz, P., Müller, M. G., Maier, M., Stoneman, S., Tomić, T., vonBargen, I., Schuster, M. J., Steidle, F., Wedler, A., Stürzl, W., & Triebel, R. (2020). ARDEA—An MAV with skills for future planetary missions. *Journal of Field Robotics*, *37*(4), 515–551. https://doi.org/10.1002/rob.v37.4

Markley, F. L. (2003). Attitude Error Representations for Kalman Filtering. *Journal of Guidance, Control, and Dynamics*, *26*(2), 311–317.

Medina, D., Vilá-Valls, J., Chaumette, E., Vincent, F., & Closas, P. (2021). Cramér-rao bound for a mixture of real- and integer-valued parameter vectors and its application to the linear regression model. *Signal Processing*, *179*, 107792.

Motaghian, S., Grindrod, P. M., Allender, E. J., Stabbins, R., Cousins, C. R., Balme, M. R., & Gunn, M. D. (2019). The ExoMars-like Field Trials (EXOFiT): PanCam Emulator Multispectral Observations. *Proceedings of the Lunar and Planetary Science Conference*.

Mur-Artal, R., Montiel, J. M. M., & Tardos, J. D. (2015). ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, *31*(5), 1147–1163.

Mur-Artal, R., & Tardós, J. D. (2017). ORB-SLAM2: An open-source SLAM system for Monocular, Stereo and RGB-D Cameras. *IEEE Transactions on Robotics*, *33*(5), 1255–1262.

NASA. (2019). NASA planetary data system—Cartography and imaging sciences discipline node. https://pds-imaging.jpl.nasa.gov/

Post, M., Michalec, R., Bianco, A., Yan, X., Maio, A. D., Labourey, Q., Lacroix, S., Gancet, J., Govindaraj, S., Martinez-Gonzalez, X., Dalati, I., Domínguez, R., Wehbe, B., Fabisch, A., Röhrig, E., Souvannavong, F., Bissonnette, V., Smíšek, M., Oumer, N. W., … Márton, Z.-C. (2018). InFuse data fusion methodology for space robotics, awareness and machine learning. *Proceedings of the International Astronautical Congress*. International Astronautical Federation.

Preston, L., Grady, M., & Barber, S. (2012). CAFE—Concepts for activities in the field for exploration—TN2: The catalogue of planetary analogues. Under ESA contract: 4000104716/11/NL/AF.

Qin, T., Li, P., & Shen, S. (2018). VINS-Mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, *34*(4), 1004–1020.

Schuster, M. J., Brunner, S. G., Bussmann, K., Büttner, S., Dömel, A., Hellerer, M., Lehner, H., Lehner, P., Porges, O., Reill, J., Riedel, S., Vayugundla, M., Vodermayer, B., Bodenmüller, T., Brand, C., Friedl, W., Grixa, I. L., Hirschmüller, H., Kassecker, M., … Wedler, A. (2019). Towards autonomous planetary exploration: The Lightweight Rover Unit (LRU), its success in the SpaceBotCamp challenge, and beyond. *Journal of Intelligent and Robotic Systems*, *93*, 461–494.

Sonsalla, R., Cordes, F., Christensen, L., Roehr, T., Stark, T., Planthaber, S., Maurus, M., Mallwitz, M., & Kirchner, E. (2017). Field testing of a Cooperative Multi-Robot Sample Return Mission in Mars Analogue Environment. *Proceedings of the Symposium on Advanced Space Technologies in Robotics and Automation*.

Strobl, K. H., & Hirzinger, G. (2006). Optimal Hand-Eye Calibration. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. (pp. 4647–4653). Beijing, China.

Strobl, K. H., & Hirzinger, G. (2011). More Accurate Pinhole Camera Calibration with Imperfect Planar Target. *Proceedings of the IEEE International Conference on Computer Vision, Workshop on Challenges and Opportunities in Robot Perception* (pp. 1068–1075). IEEE.

Strobl, K. H., Sepp, W., Fuchs, S., Paredes, C., Smíšek, M., & Arbter, K. (2020). DLR CalDe and DLR CalLab. https://rmc.dlr.de/callab

Sturm, J., Engelhard, N., Endres, F., Burgard, W., & Cremers, D. (2012). A benchmark for the evaluation of RGB-D SLAM systems. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 573–580). IEEE.

Takasu, T., & Yasuda, A. (2009). Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. *Proceedings of the International Symposium on GPS/GNSS* (pp. 4–6).

Teunissen, P. J., & Montenbruck, O. (Eds.). (2017). *Springer handbook of global navigation satellite systems*. Springer International Publishing.

Vayugundla, M., Steidle, F., Smíšek, M., Schuster, M., Bussmann, K., & Wedler, A. (2018). Datasets of long range navigation experiments in a Moon analogue environment on Mount Etna. *Proceedings of the International Symposiom on Robotics*.

Wedler, A., Vayugundla, M., Lehner, H., Lehner, P., Schuster, M. J., Brunner, S. G., Stürzl, W., Dömel, A., Gmeiner, H., Vodermayer, B., Rebele, B., Grixa, I. L., Bussmann, K., Reill, J., Willberg, B., Maier, A., Meusel, P., Steidle, F., Smíšek, M., … Wilde, M. (2017). First results of the ROBEX Analogue mission campaign: Robotic deployment of seismic networks for future lunar missions. *Proceedings of the International Astronautical Congress*. International Astronautical Federation.

Wettergreen, D., Bapna, D., Maimone, M., & Thomas, G. (1999). Developing nomad for robotic exploration of the atacama desert. *Robotics and Autonomous Systems*, *26*(2–3), 127–148.

Wettergreen, D., Cabrol, N., Baskaran, V., Calderón, F., Heys, S., Jonak, D., Lüders, A., Pane, D., Smith, T., Teza, J., Tompkins, P., Villa, D., Williams, C., & Wagner, M. (2005). Second experiments in the robotic investigation of life in the Atacama Desert of Chile. *Proceedings ISAIRAS 2005 Conference* (Vol. 603, pp. 233–240).

Woods, M., Shaw, A., Tidey, E., Pham, B. V., Simon, L., Mukherji, R., Maddison, B., Cross, G., Kisdi, A., Tubby, W., Visentin, G., & Chong, G. (2014). Seeker—autonomous long-range rover navigation for remote exploration. *Journal of Field Robotics*, *31*(6), 940–968.

How to cite this article: Meyer, L., Smíšek, M., Fontan Villacampa, A., Oliva Maza, L., Medina, D., Schuster, M. J., Steidle, F., Vayugundla, M., Müller, M. G., Rebele, B., Wedler, A., & Triebel, R. (2021). The MADMAX data set for visual-inertial rover navigation on Mars. *Journal of Field Robotics*, 1–21. https://doi.org/10.1002/rob.22016